PB89-129449

Expert Systems tc Assist in Evaluation of Measurement Data

(U.S.) Environmental Protection Agency
Cincinnati, OH

Nov 88

U.S. Department of Commerce
National Technical Information Service
NTIS

PB89-129449

EPA/600/D-88/250
November 1988

EXPERT SYSTEMS TO ASSIST IN EVALUATION OF MEASUREMENT DATA

by

Daniel G. Greathouse
Risk Reduction Engineering Laboratory
U.S. Environmental Protection Agency
Cincinnati, Ohio 45268

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

RISK REDUCTION ENGINEERING LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
CINCINNATI, OHIO 45268

TECHNICAL REPORT DATA (Please read Instructions on the reverse before completing)		
1. REPORT NO. EPA/600/D-88/250	2.	3. RECIPIENT'S ACCESSION NO. PB89-129449/AS
4. TITLE AND SUBTITLE Expert Systems to Assist in Evaluation of Measurement Data		5. REPORT DATE November 1988 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Daniel G. Greathouse		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS USEPA/RREL/WMDDRD/MSWRMB/LTS 26 West Martin Luther King Drive , Cincinnati, OH 45268		10. PROGRAM ELEMENT NO. 11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS Risk Reduction Engineering Laboratory--Cincinnati, OH Office of Research and Development U.S. Environmental Protection Agency Cincinnati, OH 45268		13. TYPE OF REPORT AND PERIOD COVERED 14. SPONSORING AGENCY CODE EPA/600/14
15. SUPPLEMENTARY NOTES Daniel G. Greathouse FTS: 684-7885 COMM: 513/569-7885		

16. ABSTRACT

The Agency expends a significant proportion of its budget to measure levels of contaminants or effects of contamination in the environment. In addition it mandates that other organizations, primarily private companies, expend considerable resources to provide measurement data to the Agency. This requires that decisions be made concerning appropriate sampling methods and analytical techniques and results in huge volumes of measurement data that must be evaluated and interpreted by Agency personnel. Concerns such as extent of contamination, potential health risks due to the contamination, and likelihood of adverse health consequences by introduction of new chemicals into the market place are just a few of the decision areas based on these data. Clearly the potential health consequences and economic implications of these evaluations and interpretations are very significant. Hence it is very important that the best expertise be brought to bear on these decisions. The purpose of this paper is to review the development of expert systems by USEPA to assist in evaluation of measurement data.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS (This Report)	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	9
	20. SECURITY CLASS (This page) UNCLASSIFIED	22. PRICE A02

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

EXPERT SYSTEMS TO ASSIST IN EVALUATION OF MEASUREMENT DATA

Daniel Greathouse
Risk Reduction Engineering Laboratory
Environmental Protection Agency
26 W. Martin Luther King Drive
Cincinnati, Ohio 45268

INTRODUCTION

The Agency expends a significant proportion of its budget to measure levels of contaminants or effects of contamination in the environment. In addition it mandates that other organizations, primarily private companies, expend considerable resources to provide measurement data to the Agency. This requires that decisions be made concerning appropriate sampling methods and analytical techniques and results in huge volumes of measurement data that must be evaluated and interpreted by Agency personnel. Concerns such as extent of contamination, potential health risks due to the contamination, and likelihood of adverse health consequences by introduction of new chemicals into the market place are just a few of the decision areas based on these data. Clearly the potential health consequences and economic implications of these evaluations and interpretations are very significant. Hence it is very important that the best expertise be brought to bear on these decisions. The purpose of this paper is to review the development of expert systems by U.S. EPA to assist in evaluation of measurement data.

The expertise required to adequately perform the necessary evaluations and interpretations of measurement data depends in part on such things as the media that are measured, the matrix of compounds in the media, and the nature of the measurements. Typically, satisfactory evaluation and interpretation will require knowledge concerning the analytical methods, sampling procedures, underlying chemistry and/or biology, effects of the matrix of compounds in the samples, etc. This knowledge will be gained partly from formal education, but much will be based on extensive relevant experience.

Scarcity of qualified personnel to design measurement studies and to evaluate and interpret the resulting data is typical in the Agency. This problem is accentuated since these decisions have largely been delegated to the EPA Regional Offices and states. Expert systems are tools that have the potential to make available the expertise required to make these decisions. These systems will make it possible to multiply the talent of experts.

An expert system is an automated process which incorporates the judgement, experience, rules of thumb, and intuition used by a human specialist to emulate that specialist's problem solving ability. Generally, the knowledge of specialists is stored in a computer in the form of facts and decision rules although many, more complex representations are available. Expert systems have several

characteristics which differentiate them from conventional software. Whereas conventional computer programs rely on numeric algorithms as the basis for their operation, expert systems are characterized by an emphasis on symbolic processing, logical inferencing, and pattern matching. Most expert systems contain facilities which allow the user to obtain an explanation concerning why a question was asked and/or to obtain clarification for a particular question if the user requires more information. In addition, the user is often supplied with the reasoning employed in reaching a conclusion. Expert systems are interactive; they do not run and calculate so much as they aid thought and offer advice.

A recently increased interest in the development of expert systems has brought to market software tools known as expert system shells. Basically, a shell simplifies the development of expert systems by providing a programming environment with a built in library of functions for common tasks such as rule generation, definition of objects, and interfacing with the user. This means that the builder of the system can concentrate on the acquisition of knowledge rather than on low level programming issues. The arrival of shells in the software marketplace is in part responsible for the success of many expert system projects.

There are a myriad of shells available to today's developers. Each shell has a different set of features and capabilities. For instance, some shells allow only text to be used for user prompts and output while other shells are more versatile. Some provide easy access to external data bases. Others provide the capability to attach the computer to certain measurement devices with readings displayed and control functions available graphically on the computer monitor. Of course, the cost of a shell is usually indicative of the breadth of features although competition in this market is increasing. For some projects a low cost/capability shell will suffice while in other environments advanced features are well worth their price. In general, shell selection is an important consideration which must be made with respect to the problem domain and the desired system functionality.

REVIEW OF EXISTING SYSTEMS

Some of the systems being developed to provide assistance in the measurement area will be summarized. For more information concerning the cited expert systems see the list of contacts provided in the appendix.

Flexible Membrane Liner System (FLEX)

Flexible membrane liners are plastic sheet materials that are used as moisture/liquid barriers in a landfill. They are placed under wastes to prevent leachate from leaking into the groundwater and on top of the wastes to keep water from entering the wastes. Typical plastic formulations for these liners are polyvinyl chloride (PVC), high density polyethylene (HDPE), and chlorosulphonated polyethylene (CSPE). Before one of these materials can be used as a liner for a hazardous waste land disposal site, the Agency requires that it be tested for chemical resistance to the anticipated leachate from the landfill. A series of physical property measurements must be performed on the liner material

after exposure to the anticipated leachate for prescribed periods of time. The test protocol is known as the Method 9090 test. An expert system, known as FLEX, has been developed to assist in evaluating and interpreting these physical property measurements. The rules are based on the input of 6 flexible membrane liner specialists. In essence the system determines if the data is adequate to fit a quadratic model and if so uses the model to determine if the predicted trend (change in the physical property measurements) over time exceeds the limits prescribed by the experts. Standard statistical methods are used to examine the data and to fit the quadratic model. The system is written in Arity Prolog for application on an IBM PC/AT class micro computer. Currently the system is being field tested.

Geophysics Expert System

This system is being designed to aid Superfund managers in selection of appropriate geophysical measurement methods. Different measurement scenarios are recommended based on the user identified purpose of the measurements. For example the system proposes a measurement scheme for situations when the goal is to assess extent of contamination as part of a site assessment. In addition to identifying the purpose of the measurements the user is also requested to supply general information concerning the site. The first version of the system is currently being reviewed and development of the second is scheduled for completion by early 1989. A follow-up will be initiated when funds become available. Each successive version is being designed to address a wider array of decision scenarios. The system is being developed in Basic for application on an IBM compatible micro computer.

Quality Assurance/Quality Control

Systems are being developed to improve the quality of measurement data produced by laboratories and to document this improvement. A system is being developed by EPA Headquarters to aid in implementation of the EPA developed data quality objective procedures/standards when designing a laboratory measurement program. More specifically it will provide three levels of assistance, namely (1) define the decision, (2) specify the information and confidence levels needed to make the decision, and (3) design a sampling plan to gather the information.

Radian Corporation has developed two prototype systems to provide sampling assistance. The smaller one aids in selection of the appropriate analytical method that will yield the desired level of accuracy for the compounds to be measured. If no method will provide the desired level of accuracy the user is supplied with a warning and information concerning the most accurate method available. Each method description includes procedures, specifications of number of QC samples (blanks and spikes, replicates and duplicates) to be taken, and a reference to the EPA document from which the text was drawn.

The larger, more complex, system is designed to aid in defining quality assurance and quality control sampling requirements. The system questions the user to determine the kind of errors to be controlled, for example; sampling or laboratory errors, biases or random errors, within

or between day variations, and so on. Based on this, the system recommends types of QC samples to be sent to the laboratory, i.e.,. spikes, duplicates, replicates. The user picks recommended kinds of QC samples, one at a time, and looks at them quantitatively. Based on how much variation is acceptable, how many samples are in a set, how rapidly (number of days of sampling) drift is to be recognized, and the desired certainty of recognizing drift if it occurs within a specified time frame, the program will recommend the number of QC samples to be taken per day. Alternatively, given the other parameters, it will derive the level of certainty from the number of samples.

This list of expert systems should not be considered as an inclusive list of all systems that have been designed to assist those responsible for decisions concerning measurement processes. Instead they illustrate the types of systems that can be and are being developed in the area.

DISCUSSION

Measurement of contaminants in the environment requires knowledge concerning a wide range of issues. Some of this information is contained in textbooks, scientific literature, and government reports. Some of the general rules or principles are only known by those who have extensive experience in the field. Expert systems provide an excellent means of compiling and synthesizing this knowledge and expertise in a form that can be accessed by persons with less training and experience. It should be emphasized, however, that these systems can not replace the expert; they can, however, extend the expert's knowledge. They can not provide advice concerning issues that can not currently be resolved by an expert or knowledgeable person in the field. For example a system can not be developed that will give advice concerning a new instrumentation method that has not yet been conceived or tested. The primary application is in the more routine situations which experts can resolve in a few hours to a few days.

There are a number of measurement and instrumentation issues that could be used as the focus for expert systems development. If the goal of developing an expert system is to aid other decision makers, it is important that the effort be successful. Success is determined by three related criteria. First, the system must accurately represent/emulate the decision processes used by knowledgeable persons in the field. Second, the system must also produce sufficient increases in decision efficiency and quality to warrant the costs for development, maintenance, and user training and support. Finally, the systems must be accepted and employed by the targeted user community. The following questions should be asked before initiation of an expert system development effort to assess the potential for success:

1. Do experts/knowledgeable persons exist who are available and willing to serve as resources in formulating the decision rules?

2. Is there consensus among the experts and/or user community concerning the decision rules appropriate for the application area addressed by the system?

3. Does the system address a definable need of the targeted user community and will knowledgeable persons from the user community participate in each stage of the development process?

4. Can the knowledge/expertise in the subject area be adequately represented with available software and hardware and will the targeted user community have the necessary input information for the system to provide meaningful advice?

5. Can a definite plan be developed and resources allocated for regular maintenance of the knowledge base and software and for providing user support and training to the target user community?

6. Are decisions in the selected area made often enough to justify undertaking software development and maintenance?

If these questions are answered in the affirmative then the potential developer must explore the scope of the problem to be addressed. If the problem is too simple an expert system will not represent a high enough payoff to justify the time and money spent developing the system. If the problem is too complex the development task will quickly become overwhelming. A rough estimation of problem scope can be measured by the time it takes an expert/knowledgeable person to reach a solution. In general an expert should be able to arrive at a decision in between one to eight hours. Additionally, the decision should be based upon implicit or explicit rules and criteria rather than strictly upon common sense. Consistent responses among decision makers and over time are also desirable features.

CONCLUSION

Expert systems offer the potential of synthesizing the knowledge and thought processes of the instrumentation experts in a form that can be used by less knowledgeable persons. In essence these systems offer an excellent mechanism for technology transfer. Some expert tools have been developed to aid in measurement related decisions. There appears to be a definite need for systems to aid specifically in instrumentation decisions and in the array of measurement issues that must be considered by the Agency and the regulated community. The apparent need for these systems, as perceived by the experts and developers, should not, however, be used as the basis for initiating development of an expert system in the area. The importance of input from the targeted user community and a thorough evaluation of system costs and benefits prior to system design and code development can not be over emphasized. The targeted user community must be actively involved during each stage of system development starting with the conceptualization stage and proceeding through the final stages of testing and maintenance of the production system. Poorly designed and maintained systems and/or unused systems are not only expensive, but tend to result in a loss in faith for a technology that offers significant opportunities to improve the efficiency of decision makers and enhance the quality of their decisions.

APPENDIX

CONTACTS FOR FURTHER INFORMATION CONCERNING THE CITED EXPERT SYSTEMS

FLEX (Flexible Membrane Liner System)

Daniel G. Greathouse
Risk Reduction Engineering Laboratory
US Environmental Protection Agency
26 W Martin Luther King Drive
Cincinnati, OH 45268

Geophysics Expert System

Aldo T. Mazzella
Environmental Monitoring and Support Laboratory
US Environmental Protection Agency
P.O. Box 15027
Las Vegas, NV 89114

Data Quality Objectives Advisor

Dean Neptune
Quality Assurance and Monitoring Support
US Environmental Protection Agency
401 M Street, S.W.
Washington, DC 20460

Quality Sampling and Analysis

Lawrence Keith
Radian Corporation
P.O. Box 9948
Austin, TX 78766